Archival Report

Brain Age Prediction Reveals Aberrant Brain White Matter in Schizophrenia and Bipolar Disorder: A Multisample Diffusion Tensor Imaging Study

Siren Tønnesen, Tobias Kaufmann, Ann-Marie G. de Lange, Geneviève Richard,
Nhat Trung Doan, Dag Alnæs, Dennis van der Meer, Jaroslav Rokicki, Torgeir Moberget,
Ivan I. Maximov, Ingrid Agartz, Sofie R. Aminoff, Dani Beck, Deanna M. Barch,
Justyna Beresniewicz, Simon Cervenka, Helena Fatouros-Bergman, Alexander R. Craven,
Lena Flyckt, Tiril P. Gurholt, Unn K. Haukvik, Kenneth Hugdahl, Erik Johnsen, Erik G. Jönsson,
Karolinska Schizophrenia Project, Knut K. Kolskår, Rune Andreas Kroken, Trine V. Lagerberg,
Else-Marie Løberg, Jan Egil Nordvik, Anne-Marthe Sanders, Kristine Ulrichsen,
Ole A. Andreassen, and Lars T. Westlye

ABSTRACT

BACKGROUND: Schizophrenia (SZ) and bipolar disorder (BD) share substantial neurodevelopmental components affecting brain maturation and architecture. This necessitates a dynamic lifespan perspective in which brain aberrations are inferred from deviations from expected lifespan trajectories. We applied machine learning to diffusion tensor imaging (DTI) indices of white matter structure and organization to estimate and compare brain age between patients with SZ, patients with BD, and healthy control (HC) subjects across 10 cohorts.

METHODS: We trained 6 cross-validated models using different combinations of DTI data from 927 HC subjects (18–94 years of age) and applied the models to the test sets including 648 patients with SZ (18–66 years of age), 185 patients with BD (18–64 years of age), and 990 HC subjects (17–68 years of age), estimating the brain age for each participant. Group differences were assessed using linear models, accounting for age, sex, and scanner. A meta-analytic framework was applied to assess the heterogeneity and generalizability of the results.

RESULTS: Tenfold cross-validation revealed high accuracy for all models. Compared with HC subjects, the model including all feature sets significantly overestimated the age of patients with SZ (Cohen's $d = -0.29$) and patients with BD (Cohen's $d = 0.18$), with similar effects for the other models. The meta-analysis converged on the same findings. Fractional anisotropy–based models showed larger group differences than the models based on other DTI-derived metrics.

CONCLUSIONS: Brain age prediction based on DTI provides informative and robust proxies for brain white matter integrity. Our results further suggest that white matter aberrations in SZ and BD primarily consist of anatomically distributed deviations from expected lifespan trajectories that generalize across cohorts and scanners.

Keywords: Bipolar, Brain age, DTI, Machine learning, Psychosis, Schizophrenia

https://doi.org/10.1016/j.bpsc.2020.06.014

Schizophrenia (SZ) and bipolar spectrum disorders are severe mental disorders with partly overlapping clinical characteristics and pathophysiology. Both are highly heritable (1), with a substantial neurodevelopmental etiology (2,3). Along with evidence of accelerated age-related brain changes in adult patients with SZ (4–6), the neurodevelopmental origin supports a dynamic lifespan perspective in which genetic and biological factors interact with age-related environmental and physiological processes.

Aberrant myelination and brain wiring during adolescence has been included among the neurobiological features of severe mental disorders, and white matter (WM) aberrations have been documented prior to disease onset (7–11). Brain imaging has shown that normative WM development follows a characteristic nonlinear trajectory, with peak maturation around the third or fourth decade (12–14). Compared with healthy control (HC) subjects, adult patients with SZ or patients with bipolar disorder (BD) exhibit anatomically distributed group-level

differences in various diffusion-based indices of WM structure (15,16).

Supporting a neurodevelopmental origin, it has been demonstrated that patients with adolescent-onset SZ show WM aberrations (17) and that their developmental trajectory is altered and delayed (18) compared with age-matched, normally developing peers. Further, children and adolescents with increased symptom burden, albeit presumably at subclinical levels, were found to exhibit altered diffusion-based WM properties compared with peers with low or no symptoms of mental distress (19), highlighting a critical role of WM development in mental health in youths. To which degree group differences observed between adult patients and HC subjects accelerate during the course of the adult lifespan is unclear. The neurodegenerative account of SZ and severe mental illness is debated (20) and lacks unequivocal support from imaging studies (16,21), but some studies have suggested stronger age-related deterioration of the brain in patients compared with HC subjects (22,23).

Despite converging evidence of case-control differences both preceding and following disease onset, recent brain imaging studies have documented substantial heterogeneity within patient groups (24,25). In contrast to conventional group-level analyses, brain age prediction using machine learning on imaging features allows for brain-based phenotyping at the individual level and enables an efficient dimensionality reduction of the neuroimaging data into one or more biologically informative summary measures (26,27). The discrepancy between an individual's chronological age and predicted brain age, referred to as the brain age gap (BAG), has been found to be higher in patients with SZ (5,28,29) and in several other brain disorders (29). However, these previous studies have exclusively used brain gray matter features for brain age prediction. Thus, given the well-documented role of WM aberrations in patients with mental illness (15,30–32), brain age prediction based on diffusion imaging is clearly warranted.

In order to fill this current gap in the literature, here we compared individual BAGs between patients diagnosed with SZ or BD and HC subjects using 4 conventional metrics (fractional anisotropy [FA], mean diffusivity [MD], radial diffusivity [RD], and axial diffusivity [AD]) obtained from diffusion tensor imaging (DTI). We used an independent training set comprising 927 HC subjects 18 to 94 years of age and applied the resulting model to our test sample including patients with SZ (n = 648), patients with BD (n = 185), and HC subjects (n = 990) from 10 independent cohorts (see Methods and Materials for details). In order to specifically assess the robustness and quantify the heterogeneity of effects across cohorts, we adopted a meta-analytic statistical framework in addition to a mega-analysis across cohorts.

We trained 6 different models based on various combinations of the DTI metrics, which allowed us to compare prediction accuracy and subsequent group differences for each model. Based on converging evidence of widespread WM aberrations in patients with severe mental disorders (15), we hypothesized higher BAG in patients with SZ and patients with BD compared with HC subjects, with stronger effects in SZ compared with BD. To test the relevance of the varying spatial resolution of the feature sets, which is important to inform the discussion regarding the anatomical specificity of brain WM aberrations, we compared models including various atlas-based tracts of interest with models including only global features. Based on previous studies comparing the age prediction accuracy and clinical sensitivity between metrics (16,27,33), we hypothesized high age prediction accuracy and sensitivity to group differences but remained agnostic concerning the relative ranking of the various features.

METHODS AND MATERIALS

We combined diffusion magnetic resonance imaging (MRI) data from 2750 individuals from 11 sites/studies across 10 different scanners. Figure 1A, Figures S1 and S2, and Tables S1 and S2 summarize key demographics for each cohort. Table S3 summarizes the MRI systems and diffusion acquisition protocols.

The dataset was split into a training set and a test set. Figure S2 shows the age distribution within each of the 2 cohorts in the training set. Briefly, the training set consisted of 927 HC subjects covering the full adult lifespan (mean age 53.81 ± 18.38 years; range, 18–94 years of age). The test set comprised 990 HC subjects (mean age = 34.70 ± 11.24 years; range, 17–68 years of age), 185 patients with BD (mean age 33.12 ± 10.53 years; range, 18–64 years of age), and 648 patients with SZ (mean age 34.49 ± 11.40 years; range, 18–66 years of age).

MRI Acquisition and Processing

A summary of MRI acquisition protocol for each cohort is presented in Table S3. Imaging analyses were performed using FSL (FMRIB Software Library) (34–36). To correct for signal loss, motion, and eddy currents, all cohorts were processed using eddy (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/eddy) (37,38). Two cohorts (TOP1 and TOP2) had collected blip-up/blip-down sequences and were additionally processed using topup (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/topup) (34,39) prior to eddy. Using an integrated framework along with correction for susceptibility-induced distortions, eddy currents, and motion, eddy detects and replaces slices affected by signal loss owing to bulk motion during diffusion encoding (38).

Fitting of the diffusion tensor was done using dtifit in FSL, yielding conventional DTI metrics, including FA, MD, RD, and AD. FA, MD, RD, and AD maps were further processed using tract-based spatial statistics (40). FA volumes were skull-stripped and aligned to the FMRIB58_FA template supplied by FSL using nonlinear registration (FNIRT) (41). Next, mean FA was derived and thinned to create a mean FA skeleton, representing the center of all tracts common across subjects. We thresholded and binarized the mean FA skeleton at FA > 0.2. The procedure was repeated for MD, AD, and RD. For each individual, we calculated the mean skeleton value for each metric, as well as the mean values within 23 tracts of interest (Table S4) based on 2 probabilistic WM atlases [CBM-DTI-81 WM labels atlas and the Johns Hopkins University WM tractography atlas (42–44)]. In total, we derived 96 DTI features per individual including the mean skeleton values.

Quality Assessment

Subjects with poor image quality due to subject motion or other visible image artifacts (e.g., due to metal) were removed (n = 160; 59 HC subjects, 39 patients with SZ, 28 patients with
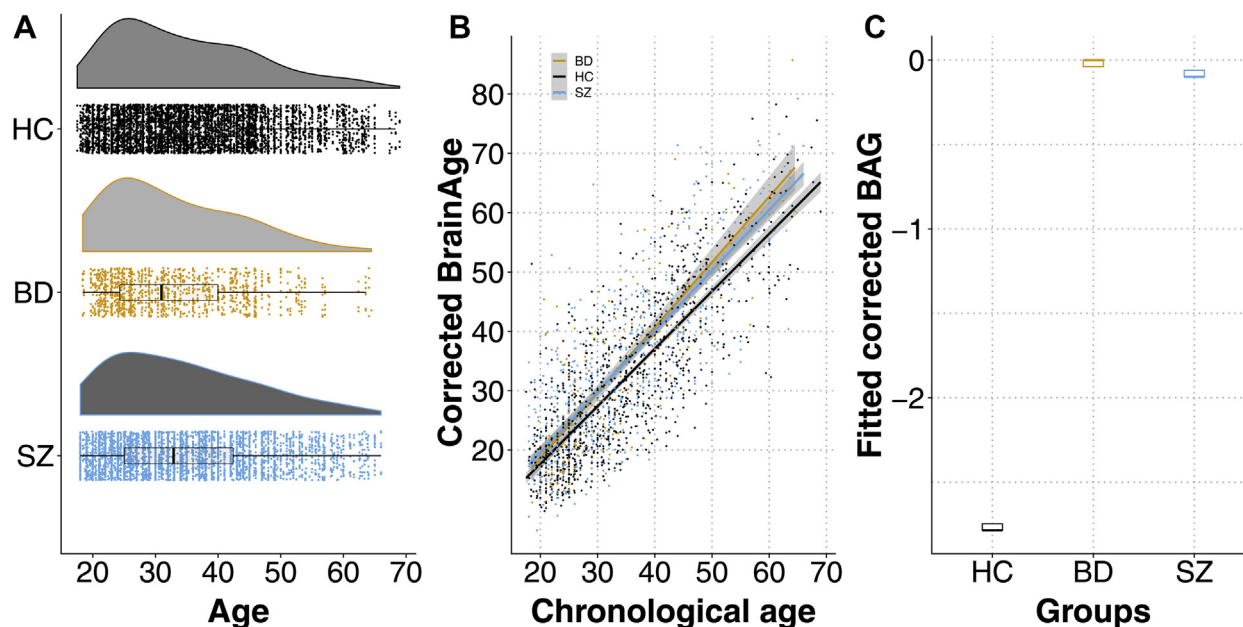
White Matter Brain Age in Severe Mental Illness

Figure 1. **(A)** Raincloud plot depicting the age distribution for each diagnostic group in the test sets. Density plots are shown on top with data points and boxplot underneath. **(B)** Predicted age (corrected for age and scanner) based on the all-features model plotted as function of chronological age. The fit lines represent the best linear fit within each group. **(C)** Boxplots plots showing the distribution of fitted corrected brain age gap (BAG) in each group. BD, bipolar disorder; HC, healthy control; SZ, schizophrenia.

BD, and 34 individuals with missing information). Demographics of the excluded participants are presented in Table S5. Additionally, we employed a multistep quality assessment (QA) procedure (16) that included maximum voxel intensity outlier count (MAXVOX) and temporal signal-to-noise-ratio (45) prior to statistical analyses. Briefly, we ran the QA iteratively, excluding participants with a QA score of 2.5 SD below the mean. In order to compute the QA score, we inverted the MAXVOX score, z-normalized both scores independently (MAXVOX and temporal signal-to-noise-ratio), and computed a summary score combing the two scores. In short, manual inspection of the flagged datasets after QA suggested adequate quality. Thus, we present results on the full dataset with supplemental results from a stringent QA [see (16) for additional information].

Brain Age Prediction

We trained 6 age prediction models. Our main model included all 96 features across all DTI metrics. To assess sensitivity for each metric separately, we trained 4 additional models based on all tracts of interest for each metric (FA, RD, MD, or AD). To test the value of including regionally specific information, we trained an additional model with only the global mean skeleton feature from all 4 metrics included.

The following pipeline for brain age prediction was identical for all 6 models: we used the xgboost framework in R version 3.3.3 (2017-03-06) (R Foundation for Statistical Computing, Vienna, Austria) (46) to build the prediction model. The number of rounds (nround), maximum depth (max_depth), and subsample were tuned and optimized using a fivefold cross-validation of the training data, with early stopping if the prediction errors did not improve for 20 rounds. Based on

previous experience, the learning rate (η) was set to $\eta = 0.01$ in order to increase transparency in the parameter selection stage. Besides the default setting, the following parameters were used in the model: nround = 1400, max_depth = 14.

Prior to implementing the model, we regressed out the main effect of scanner from the DTI features in the entire dataset while accounting for age, age^2, and sex using linear models in R. To estimate the reliability of our age prediction model, we used a 10-fold cross-validation procedure within the training sample and repeated the cross-validation step 100 times to provide a robust estimate of model prediction. Within the same procedure, we tested the performance of our trained model by predicting age in unseen subjects in the test sample. By applying the model to the test sample 100 times, we obtain both a mean estimate and an estimate of uncertainty. For each iteration, we calculated the BAG, defined as the difference between chronological and predicted age. For each individual in the test set we computed the average BAG across the 100 folds and corrected these values for main effects of scanner and a well-documented age-related bias using linear models, per previous recommendations (47). Next, based on the age- and scanner-corrected BAG, we computed a corrected brain age for each individual, and then computed the mean absolute error (MAE), root mean square error (RMSE), and correlation between corrected predicted age and chronological age as measures of model performance.

Statistical Analyses

Statistical analyses were performed using R. We tested for main effects of diagnosis using linear models with corrected BAG as dependent variable and group, sex, age, and scanner site as independent variables, and performed pairwise group

comparisons as appropriate. Using the metafor package (48) in R, we adopted a meta-analytic framework in order to assess the heterogeneity and generalizability of the results. A random-effects model was used to weigh the primary studies prior to aggregating the effect size. Effect sizes were aggregated using the estimated marginal means of the BAG from each group contrast (HC/SZ, HC/BD, and BD/SZ), accounting for age, age², and sex. For effect size estimates, we used Hedges' g. Cochran's heterogeneity statistic (Q) was used to test the homogeneity of effect sizes. A χ^2 test with $k-1$ degrees of freedom was used to examine the significance of Cochran's Q. The heterogeneity was quantified using the I^2 statistic, which is sensitive to the degree of inconsistency in results between cohorts.

RESULTS

Brain Age Predictions

Age prediction in the training set using 10-fold cross-validation revealed high correlations between chronological and predicted age for the main model including all features (r = .924; 95% confidence interval, .912–.935; MAE = 6.49; RMSE = 8.08) (Figure S3).

Figure 1B shows predicted age plotted as a function of chronological age for the unseen test set when using the full feature set, and Table 1 summarizes the prediction accuracy for all 6 models. The age prediction models generalized to HC subjects (r = .806; MAE = 6.92; RMSE = 8.46), patients with BD (r = .808; MAE = 6.85; RMSE = 8.50), and patients with SZ (r = .798; MAE = 7.11; RMSE = 8.79). While all models performed relatively well, prediction accuracy was highest for the full model, and the global mean skeleton model outperformed the region of interest–based single-metric models. Figure S4 shows the correlation matrix between all models, indicating a strong correlation between all models, with the exception of FA and AD.

Group Differences in BAG

Table 1 and Figure S5 summarize the results from the group comparisons from the 6 models, and Figure 1C shows the distributions of fitted corrected BAG within each group for the all-features model. Briefly, all models revealed significant main effects of group, with higher corrected BAG in patients with SZ and BD compared with HC subjects, with effect sizes (Cohen's d) ranging between −0.10 and 0.33. The model based on FA yielded the strongest effect size for the main group effect, although the models including MD and RD in addition to FA revealed similar patterns. The model based on AD revealed less consistent results and was the only model not showing significant group differences between patients with SZ and HC subjects.

Meta-analysis and Heterogeneity in Effects Between Cohorts

Figure 2 shows a forest plot summarizing the results from the meta-analysis for BAG computed using the full feature model. Figures S6 to S10 show the results from the other models. In short, in line with the mega-analysis, the results revealed significantly higher BAG in patients with SZ and BD

Table 1. Results From the 6 Brain Age Models

	Correlation (r) Between Chronological and Predicted Age (MAE/RMSE)[a]	Corrected BAG, Mean ± SD (95% CI)						Cohen's d			Cohen's d Mean Skeleton[c]		
		Patients With SZ	Patients With BD	HC Subjects	Pairwise Comparison	F	p[b]	HC/SZ	BD/HC	BD/SZ	SZ/HC	BD/HC	BD/SZ
All Features	.80 (6.98/8.58)	−0.09 ± 8.79 (−0.77 to 0.59)	−0.02 ± 8.53 (−1.26 to 1.22)	−2.77 ± 8.00 (−3.27 to −2.27)	HC<SZ, HC<BD	23.01	<.001	−0.29	0.18	−0.03	–	–	–
Mean Skeleton	.76 (7.91/9.75)	0.56 ± 10.04 (−0.21 to 1.34)	−0.16 ± 8.97 (−1.46 to 1.14)	−2.55 ± 9.37 (−3.13 to −1.97)	HC<SZ, HC<BD	21.57	<.001	0.30	0.14	−0.04	–	–	–
FA	.77 (7.61/9.50)	0.24 ± 9.92 (−0.52 to 1.01)	0.21 ± 9.42 (−1.16 to 1.57)	−3.23 ± 8.66 (−3.77 to −2.69)	HC<SZ, HC<BD	30.55	<.001	−0.33	0.21	<0.01	−0.36	−0.17	0.18
RD	.78 (7.65/9.32)	−0.60 ± 9.72 (−1.35 to 0.15)	−0.05 ± 8.64 (−1.30 to 1.20)	−3.45 ± 8.50 (−3.98 to −2.92)	HC<SZ, HC<BD	24.59	<.001	−0.29	0.21	−0.03	0.35	0.20	−0.14
MD	.78 (7.54/9.23)	−1.33 ± 9.40 (−2.07 to −0.61)	−0.21 ± 8.49 (−1.44 to 1.02)	−3.45 ± 8.54 (−3.98 to −2.92)	HC<SZ, HC<BD	17.27	<.001	0.22	0.20	0.06	0.24	0.17	−0.06
AD	.80 (7.15/8.80)	−2.17 ± 8.46 (−2.82 to −1.52)	−0.46 ± 8.38 (0.75 to −1.67)	−2.95 ± 8.43 (−3.48 to −2.42)	HC>BD	6.95	.001	−0.10	0.16	0.10	<0.001	0.08	0.09

AD, axial diffusivity; BAG, brain age gap; BD, bipolar disorder; CI, confidence interval; FA, fractional anisotropy; HC, healthy control; MAE, mean absolute error; MD, mean diffusivity; RD, radial diffusivity; RMSE, root mean square error; SZ, schizophrenia.

[a]Correlations, MAE, and RMSE are reported for predicted age corrected for scanner and age.

[b]α level for Bonferroni-corrected p values = .008.

[c]Cohen's d reported for the pairwise group comparisons for mean skeleton diffusion tensor imaging values.
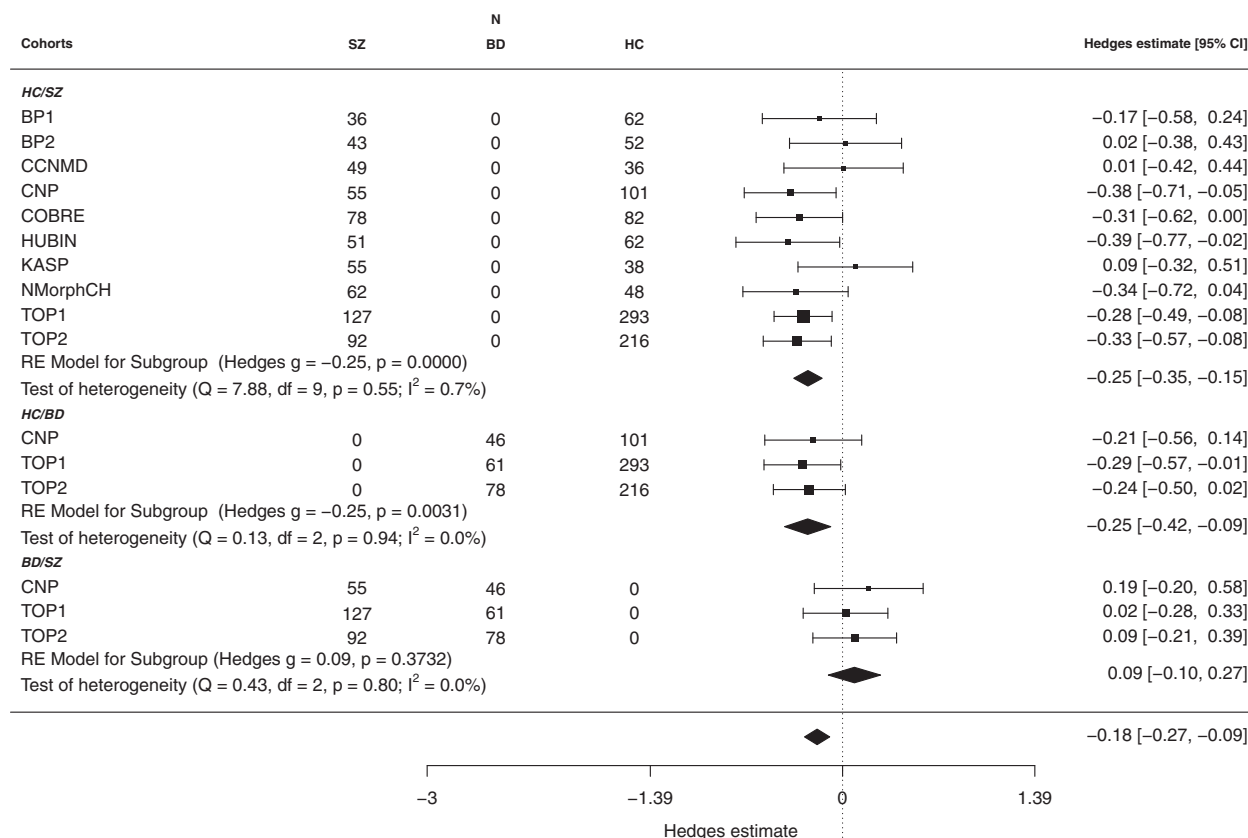
White Matter Brain Age in Severe Mental Illness

Figure 2. Forest plot summarizing the results from the meta-analytical approach for the all features model. Hedges' estimate (*g*) was used to calculate the effect size. BD, bipolar disorder; CCNMD, Conte Center for the Neuroscience of Mental Disorders; CI, confidence interval; CNP, Consortium for Neuropsychiatric Phenomics; COBRE, Center for Biomedical Research Excellence; HC, healthy control; HUBIN, Human Brain Informatics; KASP, Karolinska Schizophrenia Project; NMorphCH, Neuromorphometry by Computer Algorithm Chicago; RE, random effects model; SZ, schizophrenia; TOP, Tematisk Område Psykose/Norwegian Centre for Mental Disorders Research.

compared with HC subjects, with moderate effect sizes. The analysis did not support a group difference in BAG between BD and SZ. Whereas the effect sizes varied slightly between cohorts for the full model, the *Q* and I^2 statistics indicated low and nonsignificant heterogeneity. Figure S11 shows each cohort's contribution to the heterogeneity and influence on the result from the meta-analysis.

Quality Control

Figure S12 summarizes the results from multistep QA. Briefly, higher corrected BAG was observed in patients with SZ and BD compared with HC subjects across all levels of QA, with highly similar effect sizes.

DISCUSSION

The etiology of severe mental disorders has a substantial neurodevelopmental component, which is among other characteristics reflected in altered brain maturational trajectories during the formative years of childhood and adolescence, and as group-level differences in adult patient populations. Along with evidence of genetic and clinical overlap with several aging-related conditions, including

cardiovascular risk factors and increased mortality, the neurodevelopmental account supports the need for a dynamic lifespan perspective in the search for disease mechanisms. Here, in 10 different cohorts comprising HC subjects and patients with SZ and BD, we used machine learning to estimate brain age using DTI indices of WM structure and organization. This novel approach yielded 5 main results. First, in a large independent training set, we found high accuracy of brain age prediction across the adult lifespan using DTI features, which largely generalized to the independent test set, supporting the feasibility and sensitivity of the approach. Second, applying the model to an independent test set revealed significantly higher BAG in patients with SZ and BD compared with HC subjects. Third, follow-up meta-analysis and tests of heterogeneity suggested high consistency across independent cohorts and scanners. Fourth, brain age models based on FA showed higher sensitivity than models based on the other metrics, both alone and combined. Fifth, the reduced set of global mean skeleton features compared with a number of regional atlas-based features revealed highly converging results. We next discuss the implications of these findings in more detail.

Brain age prediction provides an informative summary measure that may serve as a proxy for brain integrity and health across normative and clinical populations. Neuroimaging-derived WM and gray matter phenotypes carry distinct biological information of brain integrity, and tissue-specific brain age models may provide higher sensitivity and specificity to relevant biological processes compared with conventional models based on gray matter features alone (27). DTI has been broadly applied in clinical neuroscience owing to its proposed sensitivity to microstructural properties of brain tissue. However, whereas previous studies have documented higher brain age in patients with severe mental disorders, these were based on gray matter models only (5,28,29). In order to test if previous findings suggesting clinical deviations from normative gray matter trajectories generalize to WM, we performed brain age prediction using different combinations of DTI metrics. In line with previous brain age prediction studies using diffusion MRI (27,49) we obtained high age prediction accuracy across most models. In accordance with previous evidence suggesting that regional DTI-based indices of brain aging reflect relatively low-dimensional and global processes (12,50), we found similar prediction accuracy for the reduced models comprising global mean skeleton values and the models including extended sets of regional features. Although brain WM aging shows some regional heterogeneity (12), these findings demonstrate that the most relevant information required for brain age prediction is captured at a global level. This conjecture is also supported by a recent twin study demonstrating that a large proportion of the estimated heritability of specific tracts is accounted for by a general factor (51).

Likewise, we found that the sensitivity to group differences was not strongly dependent on the inclusion of the full feature set. Indeed, the effect size obtained when comparing patients with SZ and HC subjects was slightly higher for the global mean skeleton model compared with the full model. These findings are in line with recent evidence of anatomically widely distributed group differences between healthy control subjects and patients with SZ (15). Interestingly, the largest effect when comparing patients with SZ and HC subjects was obtained for the FA-only model, supporting the sensitivity of FA to clinical differences in WM properties (15,16). Higher predicted brain age gap in the patient groups compared with HC subjects may indicate altered rate of brain maturation or accelerated brain aging in patients with severe mental disorders. However, our cross-sectional design does not permit us to make inference about brain development or aging per se, and previous reports of relatively age-invariant group differences in brain volumetry (21) and DTI indices (16) suggest that the reported group differences in brain age may reflect differences accumulating early in life. Unfortunately, owing to the current study design with adults only, we cannot address the maturational trajectories in the formative years. Further, the data in the training and test sets were collected using different scanners, and the absolute brain age estimates and corresponding deviance from the chronological age should be interpreted with caution and not without reference to an appropriate comparison sample. Although the application of diffusion MRI as the basis for age prediction is novel, higher gray matter brain age has been shown in several brain and mental disorders (29,52). We expand these previous findings by documenting higher DTI-based WM brain age in both SZ and BD, and although with moderate effect sizes, we show that the effects generalize relatively well across cohorts and scanners, with only minor heterogeneity in effect sizes between cohorts.

We found no significant difference in DTI-based brain age between BD and SZ, supporting previous evidence of partly overlapping clinical and biological characteristics between these 2 diagnostic categories (16,53,54). While the current results support the existence of a common set of mechanisms across disorders, future studies utilizing a broader range of imaging modalities in combination with specific genetic, clinical, cognitive, sociodemographic, and biological phenotypes may allow for the identification of specific diagnostic signatures and subgroups. However, inherent limitations associated with the classical case-control design in mental health research have recently been emphasized using neuroimaging data (24,25). In particular, the current lack of biologically informed diagnostic criteria should motivate future studies to consider alternative approaches to promote a novel clinical nosology based on both symptomatology and data-driven clustering (55), as well as brain-based and biological phenotypes cutting across diagnostic boundaries.

Our results document robust group-level deviances in WM structure manifesting as older-appearing brains in patients with severe mental disorders compared with their healthy peers. Whereas DTI-based markers are sensitive to different biological and anatomical characteristics, the current specificity does not allow for inference on the distinct neurobiological mechanisms involved. Myelin integrity and myelin packing density are among the proposed candidate mechanisms for observed changes in DTI metrics (56–58), but the specificity is low, and the current results probably reflect a combination of neurobiological processes and macroanatomical differences. Previous evidence implicated myelin-related abnormalities and neuroinflammation both in the pathophysiology of severe mental disorders and in brain aging (59–62). Future studies may benefit from the inclusion of advanced multishell diffusion MRI, allowing for stronger inference on the microstructural milieu of the brain tissue, including microstructural indices based on different diffusion scalar metrics [e.g., neurite orientation dispersion and density imaging (63,64), diffusion kurtosis imaging (65), WM tract integrity (66), and restriction spectrum imaging (67)].

In line with previous findings of widely distributed effects in well-powered studies of brain aging (12) and SZ (15), we found similar age prediction accuracy and subsequent group differences in brain age for the model including only global mean skeleton values and the model including a range of regionally informative values extracted from various atlas-based tracts and regions of interest. Although specific symptoms and clinical traits may map preferentially onto specific neuroanatomical subsystems [see, e.g., (19)], these novel results suggest that a large proportion of the variance associated with age and corresponding deviations in the patient groups is captured by primarily global brain processes, with relevance for our understanding of the anatomical heterogeneity and dimensionality of brain aging and severe mental illness.

In addition to the anatomical distribution of effects, the spatiotemporal dynamics of brain development and aging and their deviations in patients with mental disorders remain unclear.

The individual-level onset and rate of the group-level deviations from the normative WM trajectory is unknown and can only be inferred using longitudinal designs covering sensitive periods of neurodevelopment. Previous studies have shown both delayed neurodevelopment during adolescence (18) and accelerated aging in adulthood (5) in patients with severe mental disorders. Whereas these observations are not mutually exclusive, future studies should aim at disentangling the lifespan dynamics, e.g., by including individuals with a wider age range, and pursuing longitudinal designs including individuals across a wide range of functional levels and risk. The latter may be particularly pertinent to disentangle primary disease-related mechanisms and secondary factors related to the disease, including medication and lifestyle factors such as nutrition, physical activity, education, and a range of sociodemographic variables, all of which interact with key neurodevelopmental processes (68). Unfortunately, although possible effects of psychotropic drugs on the brain is a topic of great interest and importance (69–71), in common with other studies employing a cross-sectional and nonrandomized design, the current design does not allow us to make inference about the effects of medication and other clinical and lifestyle factors on brain age, which should be investigated by future and properly designed studies. Meanwhile, previous studies reporting associations with medication status in smaller samples need to be interpreted in light of the recent lack of significant associations in the largest DTI study to date (15). We did not exclude WM hyperintensities in the training or test sets, and future studies including a wider range of MRI modalities are necessary to determine the possible confounding effects of WM hyperintensities on the age prediction models and subsequent group comparisons. Whereas our procedure for correcting for scanner effects using linear models effectively removes simple main effects of scanner, even subtle differences in clinical recruitment and other participant characteristics between sites might induce interactions with site- or scanner-related variance that are very difficult to account for statistically. Additionally, future studies using different samples and approaches for brain age prediction is needed to validate and test the generalizability of the model.

In conclusion, in this multisample study including patients from 10 different cohorts, we report higher brain age in patients with SZ and BD compared with HC subjects using various DTI-based indices of WM structure and organization. In contrast to most previous studies comparing diffusion MRI metrics directly between groups, we used a multisample approach, which allowed us to specifically assess generalizability across 9 or 10 different cohorts, sites, and scanners. These results represent a highly relevant contribution to the field and an important supplement to previous reports, which have largely ignored between-sample heterogeneity and generalizability. Although the effect sizes were modest, our unique design allowed us to specifically quantify the heterogeneity and robustness of effects across cohorts and scanners, supporting that brain age prediction using diffusion MRI is a sensitive marker in the clinical neurosciences.

ARTICLE INFORMATION

From the Norwegian Centre for Mental Disorders Research (NORMENT) (ST, TK, A-MGdL, GR, NTD, DA, DvdM, JR, TM, IIM, IA, SRA, DB, TPG, UKH, EGJ, KKK, TVL, A-MS, KU, OAA, LTW), Institute of Clinical Medicine, University of Oslo; Adult Psychiatry Unit (UKH), Department of Mental Health and Addiction, Institute of Clinical Medicine, University of Oslo; Department of Psychology (A-MGdL, JR, TM, IIM, DB, KKK, A-MS, KU, LTW) and KG Jebsen Centre for Neurodevelopmental Disorders (IA, OAA, LTW), University of Oslo; Division of Mental Health and Addiction (ST, TK, A-MGdL, GR, NTD, DA, DvdM, JR, TM, IIM, IA, SRA, DB, TPG, UKH, EGJ, KKK, TVL, A-MS, KU, OAA, LTW), Oslo University Hospital; Early Intervention in Psychosis Advisory Unit for South East Norway (SRA), Division of Mental Health and Addiction, Oslo University Hospital; Bjørknes University College (DA); and Department of Psychiatric Research (IA), Diakonhjemmet Hospital, Oslo; Sunnaas Rehabilitation Hospital HF (KKK, A-MS, KU), Nesodden; Department of Biological and Medical Psychology (JB, KH, ARC, RAK), University of Bergen; Departments of Clinical Medicine (K1) (EK) and Clinical Psychology (E-ML), University of Bergen; and NORMENT (JB, ARC, KH, RAK, E-ML), Department of Biological and Medical Psychology (KH), Department of Radiology (KH), Department of Psychiatry (KH), Department of Clinical Engineering (ARC), and Department of Addiction Medicine (E-ML), Haukeland University Hospital, Bergen; and Catosenteret Rehabilitation Center (JEN), Son, Norway; Department of Psychiatry

(A-MGdL), University of Oxford, Warneford Hospital, Oxford, United Kingdom; School of Mental Health and Neuroscience (DvdM), Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, the Netherlands; Centre for Psychiatry Research (IA, SC, HF-B, LF, EGJ), Department of Clinical Neuroscience, Karolinska Institutet; and Stockholm Health Care Services (IA, SC, HF-B, LF, EGJ), Stockholm Region, Stockholm, Sweden; and the Department of Psychological and Brain Sciences (DMB), Washington University in St. Louis; and Departments of Psychiatry (DMB) and Radiology (DMB), School of Medicine, Washington University in St. Louis, St. Louis, Missouri.

Members of the Karolinska Schizophrenia Project (KaSP): Farde L[1], Flyckt L[1], Engberg G[2], Erhardt S[2], Fatouros-Bergman H[1], Cervenka S[1], Schwieler L[2], Piehl F[3], Agartz I[1,4,5], Collste K[1], Victorsson P[1], Malmqvist A[2], Hedberg M[2], Orhan F[2], Sellgren C[2]

[1]Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, & Stockholm County Council, Stockholm, Sweden

[2]Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden

[3]Neuroimmunology Unit, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

[4]NORMENT, Division of Mental Health and Addiction, University of Oslo, Oslo, Norway

[5]Department of Psychiatric Research, Diakonhjemmet Hospital, Oslo, Norway

Address correspondence to Lars T. Westlye, Ph.D., Department of Psychology, University of Oslo, P.O. Box 1094 Blindern, 0317 Oslo, Norway; E-mail: l.t.westlye@psykologi.uio.no.

REFERENCES

1. Pettersson E, Lichtenstein P, Larsson H, Song J, Agrawal A, Borglum AD, et al. (2019): Genetic influences on eight psychiatric disorders based on family data of 4 408 646 full and half-siblings, and genetic data of 333 748 cases and controls. Psychol Med 49:1166–1173.
2. O'Shea KS, McInnis MG (2016): Neurodevelopmental origins of bipolar disorder: iPSC models. Mol Cell Neurosci 73:63–83.
3. Bora E, Ozerdem A (2017): Meta-analysis of longitudinal studies of cognition in bipolar disorder: Comparison with healthy controls and schizophrenia. Psychol Med 47:2753–2766.
4. Kochunov P, Ganjgahi H, Winkler A, Kelly S, Shukla DK, Du X, et al. (2016): Heterochronicity of white matter development and aging explains regional patient control differences in schizophrenia. Hum Brain Mapp 37:4673–4688.
5. Schnack HG, van Haren NE, Nieuwenhuis M, Hulshoff Pol HE, Cahn W, Kahn RS (2016): Accelerated Brain aging in schizophrenia: A longitudinal pattern recognition study. Am J Psychiatry 173:607–616.
6. Kochunov P, Hong LE (2014): Neurodevelopmental and neurodegenerative models of schizophrenia: White matter at the center stage. Schizophr Bull 40:721–728.
7. Paus T, Keshavan M, Giedd JN (2008): Why do many psychiatric disorders emerge during adolescence? Nat Rev Neurosci 9:947–957.
8. Carletti F, Woolley JB, Bhattacharyya S, Perez-Iglesias R, Fusar Poli P, Valmaggia L, et al. (2012): Alterations in white matter evident before the onset of psychosis. Schizophr Bull 38:1170–1179.
9. Paillère Martinot ML, Lemaitre H, Artiges E, Miranda R, Goodman R, Penttilä J, et al. (2013): White-matter microstructure and gray-matter volumes in adolescents with subthreshold bipolar symptoms. Mol Psychiatry 19:462–470.
10. Barnea-Goraly N, Chang KD, Karchemskiy A, Howe ME, Reiss AL (2009): Limbic and corpus callosum aberrations in adolescents with bipolar disorder: A tract-based spatial statistics analysis. Biol Psychiatry 66:238–244.
11. Insel TR (2010): Rethinking schizophrenia. Nature 468:187–193.
12. Westlye LT, Walhovd KB, Dale AM, Bjornerud A, Due-Tonnessen P, Engvig A, et al. (2010): Life-span changes of the human brain white matter: Diffusion tensor imaging (DTI) and volumetry. Cereb Cortex 20:2055–2068.
13. Rathee R, Rallabandi VP, Roy PK (2016): Age-related differences in white matter integrity in healthy human brain: Evidence from structural MRI and diffusion tensor imaging. Magn Reson Insights 9:9–20.
14. Kochunov P, Williamson DE, Lancaster J, Fox P, Cornell J, Blangero J, et al. (2012): Fractional anisotropy of water diffusion in cerebral white matter across the lifespan. Neurobiol Aging 33:9–20.
15. Kelly S, Jahanshad N, Zalesky A, Kochunov P, Agartz I, Alloza C, et al. (2018): Widespread white matter microstructural differences in schizophrenia across 4322 individuals: Results from the ENIGMA Schizophrenia DTI Working Group. Mol Psychiatry 23:1261–1269.
16. Tonnesen S, Kaufmann T, Doan NT, Alnaes D, Cordova-Palomera A, Meer DV, et al. (2018): White matter aberrations and age-related trajectories in patients with schizophrenia and bipolar disorder revealed by diffusion tensor imaging. Sci Rep 8:14129.
17. Douaud G, Smith S, Jenkinson M, Behrens T, Johansen-Berg H, Vickers J, et al. (2007): Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia. Brain 130:2375–2386.
18. Douaud G, Mackay C, Andersson J, James S, Quested D, Ray MK, et al. (2009): Schizophrenia delays and alters maturation of the brain in adolescence. Brain 132:2437–2448.
19. Alnaes D, Kaufmann T, Doan NT, Cordova-Palomera A, Wang Y, Bettella F, et al. (2018): Association of heritable cognitive ability and psychopathology with white matter properties in children and adolescents. JAMA Psychiatry 75:287–295.
20. Rund BR (2009): Is schizophrenia a neurodegenerative disorder? Nord J Psychiatry 63:196–201.
21. Moberget T, Doan NT, Alnaes D, Kaufmann T, Cordova-Palomera A, Lagerberg TV, et al. (2018): Cerebellar volume and cerebellocerebral structural covariance in schizophrenia: A multisite mega-analysis of 983 patients and 1349 healthy controls. Mol Psychiatry 23:1512–1520.
22. Cropley VL, Klauser P, Lenroot RK, Bruggemann J, Sundram S, Bousman C, et al. (2017): Accelerated gray and white matter deterioration with age in schizophrenia. Am J Psychiatry 174:286–295.
23. Kochunov P, Glahn DC, Rowland LM, Olvera RL, Winkler A, Yang YH, et al. (2013): Testing the hypothesis of accelerated cerebral white matter aging in schizophrenia and major depression. Biol Psychiatry 73:482–491.
24. Wolfers T, Doan NT, Kaufmann T, Alnaes D, Moberget T, Agartz I, et al. (2018): Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. JAMA Psychiatry 75:1146–1155.
25. Alnaes D, Kaufmann T, van der Meer D, Cordova-Palomera A, Rokicki J, Moberget T, et al. (2019): Brain heterogeneity in schizophrenia and its association with polygenic risk. JAMA Psychiatry 76:739–748.
26. Cole JH, Marioni RE, Harris SE, Deary IJ (2019): Brain age and other bodily 'ages': Implications for neuropsychiatry. Mol Psychiatry 24:266–281.
27. Richard G, Kolskar K, Sanders AM, Kaufmann T, Petersen A, Doan NT, et al. (2018): Assessing distinct patterns of cognitive aging using tissue-specific brain age prediction based on diffusion tensor imaging and brain morphometry. PeerJ 6:e5908.
28. Koutsouleris N, Davatzikos C, Borgwardt S, Gaser C, Bottlender R, Frodl T, et al. (2014): Accelerated brain aging in schizophrenia and beyond: A neuroanatomical marker of psychiatric disorders. Schizophr Bull 40:1140–1153.
29. Kaufmann T, van der Meer D, Doan NT, Schwarz E, Lund MJ, Agartz I, et al. (2019): Common brain disorders are associated with heritable patterns of apparent aging of the brain. Nat Neurosci 22:1617–1623.

30. Klauser P, Baker ST, Cropley VL, Bousman C, Fornito A, Cocchi L, et al. (2017): White matter disruptions in schizophrenia are spatially widespread and topologically converge on brain network hubs. Schizophr Bull 43:425–435.

31. Ellison-Wright I, Bullmore E (2009): Meta-analysis of diffusion tensor imaging studies in schizophrenia. Schizophr Res 108:3–10.

32. Patel S, Mahon K, Wellington R, Zhang J, Chaplin W, Szeszko PR (2011): A meta-analysis of diffusion tensor imaging studies of the corpus callosum in schizophrenia. Schizophr Res 129:149–155.

33. de Lange AM, Barth C, Kaufmann T, Maximov II, Van der Meer D, Agartz I, et al. (2019): Cumulative estrogen exposure, APOE genotype, and women's brain aging - a population-based neuroimaging study. bioRxiv. https://www.biorxiv.org/content/10.1101/826123v1.full.

34. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, et al. (2004): Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23(Suppl 1):S208–S219.

35. Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, et al. (2009): Bayesian analysis of neuroimaging data in FSL. NeuroImage 45:S173–S186.

36. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012). FSL. NeuroImage 62:782–790.

37. Andersson JL, Sotiropoulos SN (2016): An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. NeuroImage 125:1063–1078.

38. Andersson JL, Graham MS, Zsoldos E, Sotiropoulos SN (2016): Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. NeuroImage 141:556–572.

39. Andersson JL, Skare S, Ashburner J (2003): How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. NeuroImage 20:870–888.

40. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, et al. (2006): Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. NeuroImage 31:1487–1505.

41. Jenkinson M, Bannister P, Brady M, Smith S (2002): Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage 17:825–841.

42. Hua K, Zhang J, Wakana S, Jiang H, Li X, Reich DS, et al. (2008): Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. NeuroImage 39:336–347.

43. Mori S, Wakana S, Nagae-Poetscher L, van Zijl P (2005): MRI Atlas of Human White Matter. Amsterdam, the Netherlands: Elsevier.

44. Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL, et al. (2007): Reproducibility of quantitative tractography methods applied to cerebral white matter. NeuroImage 36:630–644.

45. Roalf DR, Quarmley M, Elliott MA, Satterthwaite TD, Vandekar SN, Ruparel K, et al. (2016): The impact of quality assurance assessment on diffusion tensor imaging outcomes in a large-scale population-based cohort. NeuroImage 125:903–919.

46. Chen T, He T, Benesty M, Khotilovich V, Tang Y (2015): Xgboost: Extreme gradient boosting. R package version 04-2.1-4.

47. de Lange AG, Cole JH (2020): Commentary: Correction procedures in brain-age prediction. NeuroImage Clin 26:102229.

48. Viechtbauer W (2010): Conducting meta-analyses in R with the metafor package. J Stat Soft 36:3.

49. Mwangi B, Hasan KM, Soares JC (2013): Prediction of individual subject's age across the human lifespan using diffusion tensor imaging: A machine learning approach. NeuroImage 75:58–67.

50. Penke L, Maniega SM, Murray C, Gow AJ, Valdés Hernández MC, Clayden JD, et al. (2010): A general factor of brain white matter integrity predicts information processing speed in healthy older people. J Neurosci 30:7569.

51. Gustavson DE, Hatton SN, Elman JA, Panizzon MS, Franz CE, Hagler DJ Jr, et al. (2019): Predominantly global genetic influences on individual white matter tract microstructure. NeuroImage 184:871–880.

52. Kuhn T, Kaufmann T, Doan NT, Westlye LT, Jones J, Nunez RA, et al. (2018): An augmented aging process in brain white matter in HIV. Hum Brain Mapp 39:2532–2540.

53. Owen MJ, Craddock N, Jablensky A (2007): The genetic deconstruction of psychosis. Schizophr Bull 33:905–911.

54. Hill SK, Reilly JL, Harris MS, Rosen C, Marvin RW, Deleon O, et al. (2009): A comparison of neuropsychological dysfunction in first-episode psychosis patients with unipolar depression, bipolar disorder, and schizophrenia. Schizophr Res 113:167–175.

55. Maglanoc LA, Landrø NI, Jonassen R, Kaufmann T, Córdova-Palomera A, Hilland E, et al. (2019): Data-driven clustering reveals a link between symptoms and functional brain connectivity in depression. Biol Psychiatry Cogn Neurosci Neuroimaging 4:16–26.

56. Song SK, Sun SW, Ju WK, Lin SJ, Cross AH, Neufeld AH (2003): Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. NeuroImage 20:1714–1722.

57. Song SK, Yoshino J, Le TQ, Lin SJ, Sun SW, Cross AH, et al. (2005): Demyelination increases radial diffusivity in corpus callosum of mouse brain. NeuroImage 26:132–140.

58. Checa A, Malmqvist A, Flyckt L, Schwieler L, Samuelsson M, Skogh E, et al. (2018): Cerebrospinal fluid levels of sphingolipids associate with disease severity in first episode psychosis patients. Schizophr Res 199:438–441.

59. Hakak Y, Walker JR, Li C, Wong WH, Davis KL, Buxbaum JD, et al. (2001): Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. Proc Natl Acad Sci U S A 98:4746–4751.

60. Stedehouder J, Kushner SA (2017): Myelination of parvalbumin interneurons: A parsimonious locus of pathophysiological convergence in schizophrenia. Mol Psychiatry 22:4–12.

61. van Kesteren CF, Gremmels H, de Witte LD, Hol EM, Van Gool AR, Falkai PG, et al. (2017): Immune involvement in the pathogenesis of schizophrenia: A meta-analysis on postmortem brain studies. Transl Psychiatry 7:e1075.

62. van den Ameele S, Fuchs D, Coppens V, de Boer P, Timmers M, Sabbe B, et al. (2018): Markers of inflammation and monoamine metabolism indicate accelerated aging in bipolar disorder. Front Psychiatry 9:250.

63. Nazeri A, Mulsant BH, Rajji TK, Levesque ML, Pipitone J, Stefanik L, et al. (2017): Gray matter neuritic microstructure deficits in schizophrenia and bipolar disorder. Biol Psychiatry 82:726–736.

64. Rae CL, Davies G, Garfinkel SN, Gabel MC, Dowell NG, Cercignani M, et al. (2017): Deficits in neurite density underlie white matter structure abnormalities in first-episode psychosis. Biol Psychiatry 82:716–725.

65. Jensen JH, Helpern JA, Ramani A, Lu H, Kaczynski K (2005): Diffusional kurtosis imaging: The quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. Magn Reson Med 53:1432–1440.

66. Fieremans E, Jensen JH, Helpern JA (2011): White matter characterization with diffusional kurtosis imaging. NeuroImage 58:177–188.

67. White NS, Leergaard TB, D'Arceuil H, Bjaalie JG, Dale AM (2013): Probing tissue microstructure with restriction spectrum imaging: Histological and theoretical validation. Hum Brain Mapp 34:327–346.

68. Alnæs D, Kaufmann T, Marquand AF, Smith SM, Westlye LT (2020): Patterns of socio-cognitive stratification and perinatal risk in the child brain. Proc Natl Acad Sci U S A 117:12419–12427.

69. Kuroki N, Kubicki M, Nestor PG, Salisbury DF, Park HJ, Levitt JJ, et al. (2006): Fornix integrity and hippocampal volume in male schizophrenic patients. Biol Psychiatry 60:22–31.

70. Cheung V, Cheung C, McAlonan GM, Deng Y, Wong JG, Yip L, et al. (2008): A diffusion tensor imaging study of structural dysconnectivity in never-medicated, first-episode schizophrenia. Psychol Med 38:877–885.

71. Bollettini I, Barberi I, Poletti S, Radaelli D, Pirovano A, Lorenzi C, et al. (2015): Sterol regulatory element binding transcription factor-1 gene variation and medication load influence white matter structure in schizophrenia. Neuropsychobiology 71:112–119.